OCR BASED MOBILE APPLICATION

KT VIMAL

MES20MCA-2024

PRODUCT OWNER: MR SYED FEROZE AHAMED M

TABLE OF CONTENTS

INTRODUCTION

OCR stands for Optical Character Recognition. It is a technology that recognizes text within a digital image. It is commonly used to recognize text in scanned documents and images. OCR software is able to go through documents and make the contents machine-readable. So that they can be worked with in an electronic format. It can be used for many different documents and allows many tasks to be automated.

The aim of this project is make text data to digital data. Thus the data can be documented. So that the retrieved data can be used in several places

This application uses OCR in a way so that documentation can be done in a simple way. When we want to document something we type the content into the document using a keyboard which will take a lot of time. So to reduce the time needed and to make the documentation easy we use the OCR.

The text-to-speech and speech-to-text are also used so that the details we speak are turned to digital data and the digitalized data is read to the user. We can select the documents which are read to the user using text-to-speech. The file is saved in docx format in the mobile device.

MODULES

❖**USER**

1. Registration
2. Login
3. Scan files
4. Text to Speech
5. Speech to Text
6. Save files
7. View files
8. View user history

METHODOLOGY

In this project the OCR uses Amazon Textract.

Amazon Textract is a machine learning (ML) service that automatically extracts text, handwriting, and data from scanned documents. It goes beyond simple optical character recognition (OCR) to identify, understand, and extract data from forms and tables. Textract can extract the data in minutes instead of hours or days.

Amazon Textract is a AWS(Amazon Web Services) AL services. AWS pre-trained AI services provide ready-made intelligence for your applications and workflows. Because we use the same deep learning technology that powers Amazon.com, you get quality and accuracy from continuously learning APIs. And best of all, AI services on AWS don’t require ML experience.

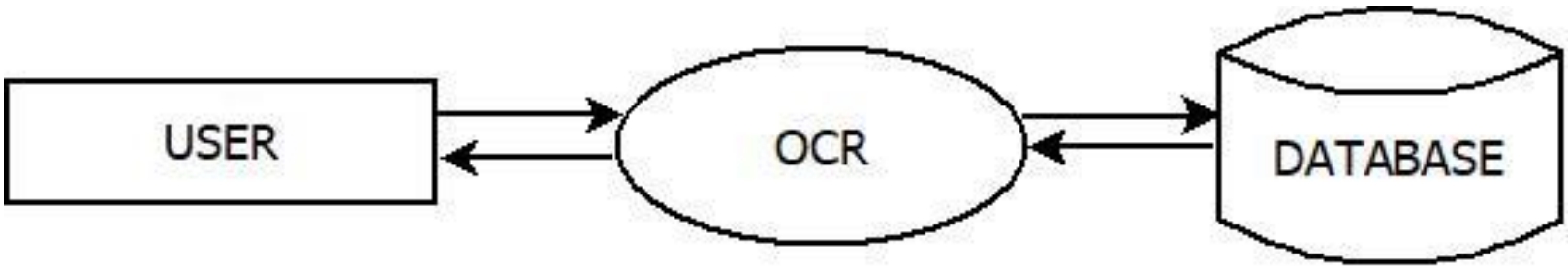
METHODOLOGY

Amazon Textract makes it easy to add document text detection and analysis to your applications. Using Amazon Textract customers can:

- Detect typed and handwritten text in a variety of documents, including financial reports, medical records, and tax forms.
- Extract text, forms, and tables from documents with structured data, using the Amazon Textract Document Analysis API.
- Process invoices and receipts with the AnalyzeExpense API.
- Process ID documents such as drivers licenses and passports issued by U.S. government, using the AnalyzeID API.

DATA FLOW DIAGRAM

LEVEL 0



LEVEL 1.1

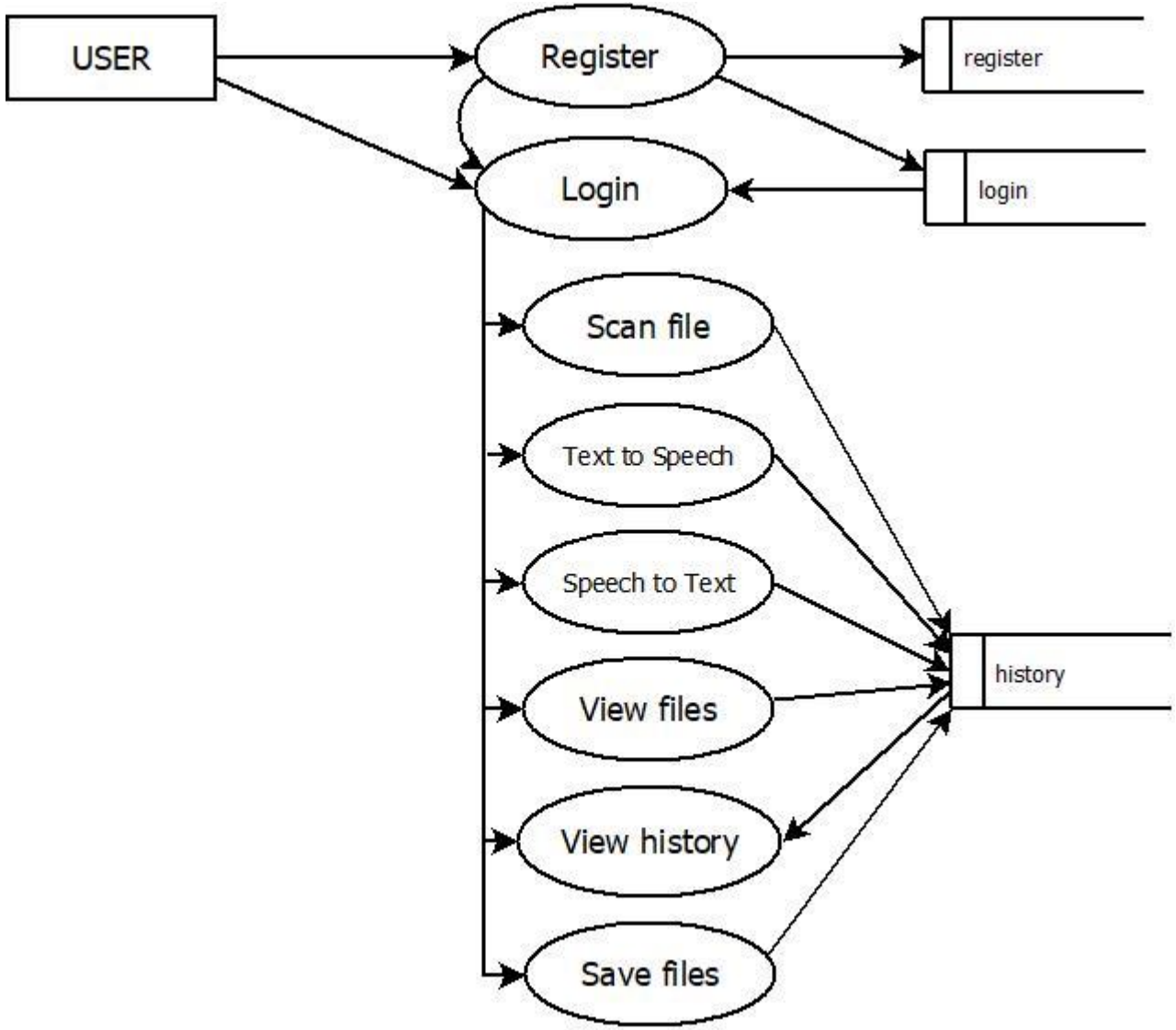

USER
Register
register
Login
login
Scan file
Text to Speech
Speech to Text
history
View files
View history
Save files

TABLE DESIGN

user

	Field	Type	Comment
	id	int(11) NOT NULL	
	lid	int(11) NULL	
	fname	varchar(54) NULL	
	lname	varchar(45) NULL	
	place	varchar(45) NULL	
	phone	varchar(45) NULL	
	email	varchar(54) NULL	

login

	Field	Type	Comment
	id	int(11) NOT NULL	
	username	varchar(45) NULL	
	password	varchar(45) NULL	
	usertype	varchar(45) NULL	

history

	Field	Type	Comment
	id	int(11) NOT NULL	
	uid	int(11) NULL	
	file	varchar(54) NULL	
	date	varchar(45) NULL	

DEVELOPING ENVIRONMENT

Hardware Requirements

- Processor - Intel x86
- Speed - 1.1 GHz
- RAM - 4 GB (min)
- Hard Disk - 50 GB

Software Requirements

- Operating System - Windows 7 or Above ,Android
- Frontend - Android
- Backend – Python ,MySQL
- Platform used - PyCharm, Android Studio ,SQLyog
- Web Browser - Google Chrome, Fire fox, Microsoft Edge
- Frame work - Flask

RESULTS & DISCUSSIONS

The text in the image is extracted by the OCR and the text is converted to digital text.

The OCR can be improved by giving it quality images so that the OCR can find the text easily.

The Text-to-Speech & Speech-to-Text gives the outputs and it can be stored in the device.

The data we retrieve using this app can be stored in ‘docx’ format.

FUTURE ENHANCEMENTS

OCR is finally moving away from just seeing and matching. Driven by deep learning, it’s entering a new phase where it first recognizes scanned text, then makes meaning of it. The competitive edge will be given to the software that provides the most powerful information extraction and highest-quality insights.

Now OCR can only recognition characters, it can be modified to identify shapes ,flowcharts ,diagrams ,graphs ,etc in the future.

PRODUCT BACKLOG

User Story ID	Priority <High/Medium/Low>	Size (Hours)	Sprint <#>	Status <Planned/In progress/Completed>	Release Date	Release Goal
1	Medium	2	1	Completed	28/12/2021	Table design
2	High	3		Completed	31/12/2021	Form design
3	High	5		Completed	08/01/2022	Basic coding
4	High	5	2	Completed	16/01/2022	Scan files
5	Medium	5		Completed	22/01/2022	Text –to-speech
6	High	5	3	Completed	27/01/2022	Speech-to-text
7	high	5		Completed	05/02/2022	Save files
8	Medium	5	4	Completed	10/02/2022	Testing data
9	High	5		Completed	19/02/2022	Output generation

USER STORIES

UserStoryID	As a <type of user>	I want to	So that I can
1	User	Login	login successful with correct username and password
2	User	Registration	user's can register with this app
3	User	Scan files	Images with text are taken and scanned using OCR
4	User	Text to Speech	The text is converted to speech
5	User	Speech to Text	The words the user says are converted to text
6	User	View files	Files are viewed from the phone storage
7	User	View history	View history of user in app
8	User	Save files	Save files in docx in storage

PROJECT PLAN

User Story ID	Task Name	Start Date	End Date	Days	Status
1	Sprint 1	26/12/2021	27/12/2021	2	Completed
2		29/12/2021	31/12/2021	3	Completed
3		03/12/2021	08/01/2022	5	Completed
4	Sprint 2	09/01/2022	16/01/2022	8	Completed
5		18/01/2022	22/01/2022	5	Completed
6	Sprint 3	23/01/2022	27/01/2022	5	Completed
7		30/01/2022	05/02/2022	7	Completed
8	Sprint 4	06/02/2022	10/02/2022	5	Completed
9		16/02/2022	19/02/2022	4	Completed

SPRINT BACKLOG PLAN

Backlog Item	Status & completion date	Original estimate in hours	Day1	Day2	Day3	Day4	Day5	Day6	Day7	Day8	Day9	Day10	Day11	Day12	Day13	Day14
User story #1,#2,#3		hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs
Table design	27/12/2021	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Form design	31/12/2021	3	0	0	0	1	1	1	0	0	0	0	0	0	0	0
Coding	08/01/2021	5	0	0	0	0	0	0	0	0	0	1	1	1	1	1
User story #4,#5																
Scan files	16/01/2022	5	1	1	0	1	0	1	0	1	0	0	0	0	0	0
Text–to-speech	22/01/2022	5	0	0	0	0	0	0	0	0	0	1	1	1	1	1
User story #6,#7																
Speech-to-text	27/01/2022	5	1	1	1	1	1	0	0	0	0	0	0	0	0	0
Save files	05/02/2022	5	0	0	0	0	0	0	0	1	0	1	1	1	0	1
User story #8,#9																
Testing data	10/02/2022	5	1	1	1	1	1	0	0	0	0	0	0	0	0	0
Output generation	19/02/2022	5	0	0	0	0	0	0	0	0	0	2	1	1	1	1
Total		40	4	4	2	4	3	2	0	2	0	5	4	4	3	4

SPRINT ACTUAL

Backlog Item	Status & completion date	Original estimate in hours	Day1	Day2	Day3	Day4	Day5	Day6	Day7	Day8	Day9	Day10	Day11	Day12	Day13	Day14
User story #1,#2,#3		hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs
Table design	27/12/2021	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Form design	31/12/2021	3	0	0	0	1	1	1	0	0	0	0	0	0	0	0
Coding	08/01/2021	5	0	0	0	0	0	0	0	0	0	1	1	1	1	1
User story #4,#5																
Scan files	16/01/2022	5	1	1	0	1	0	1	0	1	0	0	0	0	0	0
Text–to-speech	22/01/2022	5	0	0	0	0	0	0	0	0	0	1	1	1	1	1
User story #6,#7																
Speech-to-text	27/01/2022	5	1	1	1	1	1	0	0	0	0	0	0	0	0	0
Save files	05/02/2022	5	0	0	0	0	0	0	0	1	0	1	1	1	0	1
User story #8,#9																
Testing data	10/02/2022	5	1	1	1	1	1	0	0	0	0	0	0	0	0	0
Output generation	19/02/2022	5	0	0	0	0	0	0	0	0	0	2	1	1	1	1
Total		40	4	4	2	4	3	2	0	2	0	5	4	4	3	4

THANK YOU